Meeting report

Aging in the postgenomic era: simple or complex? Trudy FC Mackay

Address: Department of Genetics, North Carolina State University, Raleigh, NC 27695, USA. E-mail: trudy_mackay@ncsu.edu

Published: 16 October 2000

Genome Biology 2000, I(4):reports4018.1-4018.6

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2000/1/4/reports/4018

© Genome**Biology**.com (Print ISSN 1465-6906; Online ISSN 1465-6914)

A report on the Third Genetic Effects on Aging Meeting, The lackson Laboratory, Bar Harbor, Maine, August 4-8, 2000.

As the world population rapidly grows older, understanding the genetic and environmental factors that contribute to 'healthy' aging is becoming one of the most important social and health challenges facing the next half-century. Clearly, there is great variation in total life span and rate of aging both between and within species. Longevity is a typical quantitative trait: variation in life span within natural populations is caused by both genetic and environmental effects, with estimates of heritability (the fraction of the observed variation that is attributable to genetic variation) ranging from 10-30%.

As for all quantitative traits, the questions that need to be answered before we can claim a comprehensive understanding of the genetics of longevity are: At what loci can mutations affecting longevity arise? What loci affect naturally occurring variation in longevity - between strains, within a population, between populations, and between species? What are the allelic effects at these loci? That is, what are the homozygous and heterozygous effects, interactions with other loci (epistasis), and effects on other traits (pleiotropy)? How are these effects modulated by changes in the environment, and in males and females? And finally, what are the gene frequencies of these alleles in natural populations? One consensus that emerged from the Third Genetic Effects on Aging Meeting is that it is rather easier to formulate the questions than it is to determine their answers!

The genetics of longevity: simple or complex?

In his keynote address, George Martin (University of Washington at Seattle, USA) reviewed the past 50 years of research on mechanisms of senescence. This history is one of enthusiastic and often acrimonious debate, on several fronts.

On the one hand, viewpoints tended to polarize between 'complificationists', who perceive aging as genetically complicated, and 'simplificationists', who seek to find one or a few major mechanisms that will explain why we age. But even within these groups there are di-, tri- and tetrachotomies. Evolutionary biologists have sought to formulate theories as to why organisms age, which have usually proposed complex gene action. Because the intensity of natural selection declines with age, alleles with deleterious late-age specific effects but whose effects are neutral (in the mutation accumulation theory) or beneficial (in the antagonistic pleiotropy theory) early in life will achieve intermediate frequencies in populations. That there is substantial genetic variation for longevity has been documented by rapid response to artificial selection for postponed senescence in Drosophila. Physiological and genetic studies of strains that have undergone such selection have shown that long-lived lines of flies have higher rates of reproduction late in life (not surprisingly, as this was the trait selected) and can have reduced early fecundity and increased resistance to starvation, dessication, heat and oxidative stress compared to nonselected controls. Unfortunately, the same suite of correlated traits is not observed for lines originating from different starting populations, which argues against a single common mechanism. Further, the very definition of senescence as 'an increase in mortality with advancing age' has been challenged by the results of recent demographic studies in flies and humans showing that the rate of mortality does not decline among the oldest old.

The counterpoint to the 'complicated' viewpoint comes from studies showing that mutations at single genes can have major effects on increased longevity in yeast, the nematode *Caenorhabditis elegans*, *Drosophila* and mice, and that these genes may participate in a few key metabolic pathways. Attractive hypotheses for a 'unitary' basis for cellular aging include telomere shortening, accumulation of somatic mutations in mitochondria, oxidative stress and cellular damage by free radicals, errors in DNA replication and repair, and

altered rates and accuracy of protein synthesis. Further, the one simple factor that leads to increased longevity in mammals is dietary restriction.

George Martin gauged the temper of the participants with respect to the simple versus complex issue at the beginning of the meeting by asking for a vote on a scale of 1 (simple) to 10 (complex). The mean response was 5.4, but the range was from 2 to 10, showing that we were a representative sample of this particular research community. Of course, as with all debates, the most likely answer is that many viewpoints are to some extent correct, and the real question becomes one of determining all of the potential genetic mechanisms and the extent to which each contributes to genetic variation in aging in nature.

Several complementary approaches to identifying genes affecting longevity were discussed and presented at this meeting: mapping quantitative trait loci (QTLs) affecting naturally occurring variation in life span; analysis of mutations at loci associated with increased life span; and gene expression studies. The net result of this work is to identify a number of candidate genes to target for genotype-phenotype association studies in natural populations.

QTL mapping (see Box 1) is deceptively simple, and a few caveats were presented at the meeting. First, the statistical test may have to be sophisticated - a t-test is not enough. Gary Churchill (The Jackson Laboratory, USA) argued convincingly for adopting simple single-marker regression methods rather than more sophisticated maximum likelihood methods. Linear models such as marker regressions are flexible and allow the incorporation of multiple markers into the model, as well as marker interactions. Further, QTL mapping requires doing lots of tests for associations of marker status with trait phenotype. Relying on the 5% significance threshold given in the statistics textbook for your test statistic will mean that 5% of your results will be false positives. With wide availability of high speed computers, this problem is overcome by randomly permuting the trait and genotype data 1000 times or more and creating your own distribution of the test statistic under the null hypothesis. This technique was popularized by Churchill and his colleagues, and is now the gold standard of proof for a significant QTL. However, linear models do not work well with lots of missing data, such as occur if not all animals have genotype data at all marker loci, or one is interested in two locus interactions and all possible genotypes are not present in the mapping population. Churchill presented his new and powerful method for dealing with these problems in a Bayesian statistical framework.

Those are just the statistical issues. It goes without saying that one can only map the QTLs that are segregating in the parental strains used, and that a whole different set of QTLs could be found with other starting stocks. Richard Miller (University of Michigan, USA) and David Harrison (The Jackson Laboratory, USA) showed that using a four-way cross of different inbred strains, even incorporating an interspecific cross, can greatly increase the genetic variation in the mapping population and increase the chances of detecting QTLs. Another technique for capturing relevant alleles at many loci affecting longevity is to use parent strains that have undergone selection for the trait, as shown by James Curtsinger (University of Minnesota, USA). Robert Smookler Reis (University of Arkansas, USA) used a modification of a half-sib design, extended to QTL mapping, whereby each of three different strains of C. elegans were crossed to a fourth, common, parent.

The good news reported at this meeting is that QTLs for longevity have been mapped in mice (by Miller and Harrison), flies (by my group and by Curtsinger), and worms (by Smookler Reis and Thomas Johnson, University of Colorado, USA). Brad Rikke (University of Colorado at Boulder, USA) presented data showing substantial genetic variation in physiological responses to dietary restriction among inbred strains of mice, and evidence for QTLs affecting variation in the short-term, dramatic drop in body temperature accompanying dietary restriction and for body weight loss on food restriction. In humans and other outbred populations, QTL mapping is based on marker allele sharing methods among affected relatives, combining information over a number of pedigrees. Thomas Peris (Beth Israel Deaconess Medical Center, USA) summarized early evidence for QTLs affecting the attainment of extreme old age (100 years)

Box I: QTL mapping

In model organisms such as *Drosophila* and mice, all one requires are two inbred strains in which different alleles at loci affecting variation in the trait of interest are fixed, and a polymorphic molecular marker linkage map. Then one creates a mapping population of back-cross, F2, recombinant inbred (RI) or other segregating generations derived from the parental strains, and determines the phenotype (e.g., longevity) and multilocus genotype of each of the individuals in the mapping population. At its simplest, QTL mapping involves going through the genome, one marker at a time, dividing the individuals into marker genotype classes, and doing a statistical test to determine whether there is a significant difference in mean life span between the marker genotype classes. If there is such a difference, then the QTL is linked to the marker.

from a genome scan of pedigrees collected as part of the New England Centenarian Study.

Finding QTLs is hardly novel. More interesting is that these QTLs often have very 'flexible' (one could say 'complicated') properties. Churchill gave the first hint of this complexity when he illustrated an application of his Bayesian mapping method to detecting QTLs for blood pressure in mouse back cross mapping populations. The initial genome scan revealed three significant QTLs. However, a genome scan for pair-wise interaction effects (epistasis) detected two significant interactions involving three different QTLs: the third of these was significant only in the genetic background of either of the other two QTLs, and none had main effects that were large enough to be formally significant. A similar picture of multiple epistatic QTL interactions emerged from a study of diabetes phenotypes (blood glucose levels, body weight). Further, some of the genotype by genotype interactions were environment-specific, and only apparent when the mice were fed high fat diets.

Variable effects of longevity QTLs that depend on sex, genotype and external environment may be common. Miller detected five QTLs affecting mouse life span, of which four had sex-specific effects (i.e., significantly affected either males or females). Three of these loci exhibited significant (and sex-specific) epistatic interactions, while an additional epistatic interaction was observed between a pair of loci without significant main effects. I have mapped longevity OTLs in *Drosophila* in recombinant inbred lines derived from one pair of inbred lines, neither of which was selected for longevity. Life span of each line was measured in control and stressful environmental conditions (heat and cold stress, heat shock, starvation stress, high larval density), and in both homozygous and heterozygous background genotypes. A total of 25 QTLs were found, and all were sexand/or environment-specific. Many of these interactions were 'antagonistic' in that the QTL had opposite effects in the sexes or environments. The others were conditionally neutral, in that the QTL was significant in one sex or environment but not the other(s). Many sex-specific epistatic interactions were observed that were also environment-specific. Curtsinger did not find sex-specific QTLs for Drosophila longevity in his study using lines selected for postponed senescence. This is possibly a consequence of selection for late reproduction in both males and females. It is interesting that variable allelic effects are actually predicted by evolutionary theory. The strains in which longevity QTLs were mapped are (admittedly small) samples of naturally occurring allelic variation. Alleles that are polymorphic in natural populations must either be selectively neutral; unconditionally deleterious, and maintained by a balance between mutation and directional selection; or maintained by a balance of selective forces. The latter 'balance' could be opposite effects in males and females, different environments, or different genotypes.

From these studies, it appears that genetic variation for longevity in any one environment can be accounted for by a small number of segregating loci with moderate to large effects. However, it would not be wise to generalize this result to nature. In addition to the problem of restricted genetic samples in these mapping studies, it is important to recognize that the apparent moderate to large effects of the mapped QTLs could be artefacts of small sample sizes. The minimum effect size of a QTL is contingent on the sample size of the mapping population; therefore, it is always possible that more QTLs with smaller effects could be mapped with larger numbers of individuals or recombinant inbred lines. My results also suggest that more QTLs will be found by varying the environmental conditions.

Identifying genomic regions containing longevity QTLs is only the first step in what is likely to be a long voyage towards identifying the genetic loci that correspond to QTLs. Typically, QTL mappers construct congenic introgression lines for each QTL - that is, the QTL is back crossed into one of the parental inbred strains. This is necessary both to confirm the existence of the QTL, and define its map position more precisely. Smookler Reis showed that each of four C. elegans longevity OTLs could be fine-mapped in this manner. However, the precision of recombination mapping depends on obtaining informative recombinants, the likelihood of which decreases with physical distance. Prohibitively huge sample sizes (of the order of 104 or greater) are necessary to resolve QTLs to genomic regions containing 10 or 20 genes. I showed that quantitative deficiency complementation mapping could be used in *Drosophila* in place of recombination mapping. Using this method, longevity QTLs were mapped to 50-200 kilobase pair regions, containing from 3 to 14 genes (and predicted genes). This study also showed that a single QTL usually turned out to be composed of multiple tightly linked QTLs, however. The good news is that fine mapping can, with effort, lead us to candidate genes (at least in organisms with genome sequence projects), and that those candidate genes are not always what would have been predicted based in mutagenesis studies. QTL mapping thus enables us to identify genes affecting longevity that are actually segregating in nature, to assign function to predicted genes, and to identify novel pleiotropic effects of known loci that have been characterized for other phenotypes.

Single gene mutations

Cellular aging is accompanied by a decline in metabolic activity, decreased resistance to stress, gene dysregulation, and decreased rates of protein synthesis and degradation. Some of the genetic players in these processes are being identified by studies of single gene mutations in yeast, *C. elegans* and *Drosophila*.

'Petite' yeast cells are smaller and longer-lived than wildtype cells, as a consequence of impaired mitochondrial function. Michal Jazwinski (Louisiana State University, USA) showed that defective mitochondria in petite cells induce the 'retrograde response' pathway, in which expression of multiple nuclear genes is altered. The change in expression results in changes in intermediary metabolism, including induction of gluconeogenesis. This regulation involves the *ras2* gene product, which functions in the mitochondrial-nuclear signaling pathway, and interacts with the nuclear genes *Rtg1p* and *Rtg3p*, which in turn regulate gene transcription in the nucleus.

Aging in yeast cells is also accompanied by loss of transcriptional silencing, a shift from closed to open chromatin configuration, and excess and inappropriate transcription (gene dysregulation). The sir2 locus promotes silencing at mating type loci, telomeres and rDNA; loss-of-function mutations at sir2 are associated with shortened life span and production of extra-chromosomal rDNA circles (Jazwinski; Beatrice Jegalian, Massachusetts Institute of Technology, USA). The sir2 gene product is an nicotine adenine dinucleotide (NAD)-dependent histone deacetylase, linking deacetylation activity, which is essential for gene silencing and extended life span, with the energy status of cells. Yeast life span is increased by reducing glucose levels, environmentally or genetically. Jegalian presented a model linking silencing, metabolism and aging, in which caloric restriction leads to increased production of Sir2 and NAD synthesis, leading to silenced chromatin, which in turn prevents genome instability and inappropriate gene expression. It is possible this cellular mechanism is conserved from yeast to humans: murine Sir2-alpha has the same NAD-dependent histone deacetylase enzyme activity as yeast Sir2 (Jegalian), and the human sir2 homolog may be part of the complex that induces transcriptional silencing in human cells (Olivia Cabello, Baylor College of Medicine, USA).

Mitochondrial-nuclear signaling pathways could also be implicated in regulating the life span of C. elegans (Siegfried Hekimi, McGill University, USA). identified Several 'clock' genes (clk-1, clk-2, clk-3, gro-1) were identified in a screen for viable maternal-effect mutations. Mutations in these genes result in slowed embryonic and post-embryonic development, slowed adult rhythmic behaviors, and delayed reproduction. All mutants are long-lived. A subsequent screen for clk-like mutants that do not show a maternal effect resulted in six new time warp (twp) genes, which also have increased life span. A spontaneous suppressor mutation that suppresses the twp-1 mutation is encoded by mitochondria, suggestion that twp-1 plays a role in respiration. Hekimi postulated that the long life of *clk* mutants could be attributable to slow respiration and consequent low rate of oxygen radical production. However, Jacques Vanfleteren (University of Ghent, Belgium) presented data from direct measurement of metabolic rate in synchronous cultures of C. elegans that clk-1, clk-2, clk-3 and gro-1 actually have elevated respiration rates and ATP levels.

If decreased rates of protein synthesis accompany aging, could the aging process be attenuated by up-regulating protein synthesis? Perhaps so, suggest Monica Driscoll (Rutgers State University, USA) and Alexey Ryazanov (Robert Wood Johnson Medical School, USA). The elongation factor eEF-2 is essential for the elongation step in protein synthesis. Phosphorylation of eEF-2 by eEF-2 kinase inhibits translation and down-regulates protein synthesis. The C. elegans null mutation of eEF-2 kinase (efk-1) has elevated rates of protein synthesis and turnover, delayed development and extended life span. Overexpression of eEF-2 kinase reduces life span. During caloric restriction and starvation, protein synthesis and degradation rates increase and eEF-2 kinase activity is decreased; eEF-2 kinase is therefore a component of the molecular mechanism that senses caloric restriction.

When resources are scarce, *C. elegans* undergoes a change in life style from a motile, reproducing organism to a dauer larval form that is non-reproductive and stress-resistant. Genetic and molecular analysis of mutations in genes controlling this switch that are either constitutive-dauer or dauer-defective have revealed links between longevity, body size, metabolism, stress resistance, the central nervous system and reproduction that might be universal.

Hypomorphic mutations at *daf-2* and *age-1* are associated with a doubling of life span and increased resistance to heat, ultraviolet radiation and oxidative stress. These genes have been cloned and found to encode the *C. elegans* homolog of the insulin and insulin-like growth factor receptor, and phosphatidylinositol-OH kinase, respectively. Thus, the insulin signaling pathway is implicated in the genetic control of life span. Marc Tatar (Brown University, USA) presented data showing that this pathway also affects life span in *Drosophila melanogaster*, and John Papaconstantinou (University of Texas Medical Branch, USA) and Kevin Flurkey (The Jackson Laboratory, USA) presented similar evidence for mice.

Mutations in the *Drosophila* homolog of the insulin-like receptor (InR) and the insulin-like receptor substrate (chico) yields dwarf individuals with delayed development. InR and chico mutant females are sterile, with non-vitellogenic ovaries. This ovarian phenotype resembles that of reproductive diapause in this species, in a direct parallel to the dauer larva of C. elegans, a diapause stage. Some heteroallelic combinations of mutant alleles at *InR* can extend female life span, but the effects on longevity are complex. InR mutant flies have higher concentrations of triglycerides and increased activities of the anti-oxidant enzyme, Cu-Zn superoxide dismutase (Cu-Zn Sod). In mice, life span extension by 50% or greater has been shown for two dwarfing mutations, the Snell dwarf (Pit1dw) and the Ames dwarf (Prop1df), both of which are defective in the production of several pituitary hormones, including growth hormone. Interestingly, the life span extension in the *Pit1*^{dw} homozygotes is strongly dependent on social environment and background genotype: They are sterile, fat, and have reduced levels of circulating insulin relative to control litter mates, both as young and old animals.

There is thus a recurring theme of a genetic or environmental starvation/caloric restriction response in which reproduction ceases, stress resistance increases, and life span is extended. What is the molecular link between poor nutritional conditions and reproduction? Again, clues come from studies in C. elegans. Cynthia Kenyon (University of California at San Francisco, USA) presented evidence that there are two ligands for the daf-2 gene product, one regulated by the sensory system, and the other regulated by the somatic gonad. Olfactory mutants such as tax-4, a cyclic-nucleotidegated channel mutant, are long lived, perhaps because they are unable to respond to signals conveying the quality of the environment. Gonad ablation, whether environmentally or genetically induced, also increases life span, and the germline signaling and sensory perception pathways are independent: sensory mutants live longer when their germ line is killed.

Gene expression changes in aging

The availability of complete or advanced genome sequence projects for yeast, *C. elegans*, *D. melanogaster*, mouse and human have enabled genome-wide studies of global, coordinated changes in gene expression with age or with environmental treatments such as caloric restriction.

Tomas Prolla (University of Wisconsin at Madison, USA) reported results of his groups' pioneering surveys of changes in the abundance of transcripts of 6,347 murine genes (approximately 5-10% of the total genome) with aging and caloric restriction. The profile of changes in gene expression with advancing age in brain tissue were consistent with inflammatory and oxidative stress responses and were attenuated by caloric restriction. In skeletal muscle, aging resulted in a gene expression pattern indicative of a marked stress response and lower expression of metabolic and biosynthetic genes. Most of these changes in gene expression were prevented or ameliorated by caloric restriction, which appears to retard aging in skeletal muscle by increasing protein turnover and decreasing macromolecular damage. The expression chip analyses reiterate some of the findings of single mutant analyses, which engenders confidence that this method (despite some technical caveats) will be useful in identifying novel genes and pathways that can be explored through genetic analyses. Eugenia Wang (University of Louisville, USA) is anticipating the applicability of expression array technology to human aging and has, with her colleagues, collected RNA samples from a cross-sectional Danish population of 215 centenarians, 56 nonagarians and 96 younger controls, and a Chinese population in Taiwan consisting of seven large families with five generation pedigrees, each with a centenarian matriarch or patriarch.

Whole-genome expression assays could yield an embarrassment of riches, however. Johnson reported similar studies in young and old C. elegans. He noted that 300 transcripts were up-regulated and 200 transcripts were down-regulated in old relative to young worms, including up-regulation of 31 predicted oxidative stress genes and 35 heat-shock inducible genes. Which of these many changes in transcript abundance are causal or rate-limiting? One approach is to assess the effects of over-expression of a transgenic construct of the candidate gene in question. For example, aging is accompanied by increased cellular damage from oxidative stress. If the cellular response to oxidative stress limits life span, over-expression of genes that function to reduce oxidative stress should be accompanied by an increase in life span. John Tower (University of Southern California, USA) showed that a two-fold increase in expression levels of catalase (Cat), an anti-oxidant enzyme, had a slight negative effect on life span of D. melanogaster. However, over-expression of either Cu-Zn Sod (encoded by the nucleus) or MnSod (encoded by the mitochondria) under the control of a heat-shock promoter resulted in an increase in life span that was proportional to the amount of over-expression. From these studies, one could conclude that Cat is present in excess in normal cells, but Sod is limiting. However, William Orr (Southern Methodist University, USA) conducted similar experiments, but using the native promoter to induce MnSod expression in appropriate tissues. He did not find an increase in longevity in transgenic lines over-expressing one extra copy of MnSod, suggesting that the naturally evolved level of MnSod activity in Drosophila is near the optimum required under normal conditions.

Tower presented additional results illustrating that candidate genes whose expression patterns are altered during normal aging or in laboratory lines selected for postponed senescence may not always, in isolation, cause extended longevity when over-expressed. The small *Drosophila* heat shock proteins (Hsp22) are induced over 300-fold in all cell and tissue types during aging and in Drosophila lines selected for postponed senescence, but no life span extension was observed upon a 100-fold over-expression of Hsp22. Are these results negative because the target protein is not limiting, or because concerted changes in multiple proteins are typically required before life span extension is observed? Gordon Lithgow (University of Manchester, UK) showed that pharmacological compounds that mimicked the effects of over-expressing Sod and Cat together could rescue the short-lived phenotype of C. elegans mutations that had increased sensitivity to oxidative stress. Orr indicated that the life span of Drosophila lines triply transgenic for disulfide reductase, Cat and Cu-Zn Sod could be extended by 10%.

Genotype-phenotype associations

Clearly, constructing transgenic lines that over-express each candidate gene whose expression changes with aging, and all possible combinations of genes, becomes impossibly tedious as the number of candidate genes increases. Further, these studies do not shed light on the question of whether the candidate genes contribute to naturally occurring variation in longevity, and are limited in any case to model systems amenable to mutagenesis and germ-line transformation. How is it possible to determine what genes affect variation in aging in an outbred population for which genetic manipulations are impossible, such as humans? Prospects for success rest on the population genetic concept of 'linkage disequilibrium' (LD), explained in Box 2.

LD mapping has been successful in demonstrating association between molecular variation at candidate genes affecting Drosophila sensory bristle number and phenotypic variation in bristle number (my group). There are, however, several technical challenges to overcome and issues to consider in large scale application of this method to the genetic dissection of human longevity. First, there are possibly hundreds of candidate genes to include in such a study. Assuming that the effect of each candidate gene individually could be small, that the candidate genes could interact with each other to affect life span, and that single and multiple-locus effects could be contingent on sex and (unknown) environmental circumstances, huge sample sizes will be necessary for such studies. The immediate practical limitation is the cost of rapid, accurate and cost-effective (cheap) genotyping of single nucleotide polymorphisms (SNPs) at candidate genes. Jan Vijg (University of Texas Health Science Center, USA) illustrated the utility of one such technique - twodimensional gene scanning (TDGS). This method combines the high accuracy of denaturing gradient gel electrophoresis with high throughput from extensive PCR multiplexing. A mixture of PCR amplicons covering the target DNA sequence is subjected to two-dimensional electrophoresis, with separation in the first dimension by size and in the second dimension by melting characteristics. Amplicons with mutations are obvious from their shift in position on the gel. However, recent admixture of populations with different gene frequencies of the SNP marker and mean longevity, or recent mutation at the SNP marker, can induce spurious associations between molecular markers and longevity that are not attributable to close physical linkage. Replication of associations in multiple populations will be necessary before causality is implicated, as perhaps is the case for the beneficial effect of the APOE2 and detrimental effect of the APOE4 allele on attaining centenarian status (Wang).

Simple or complex aging revisited

The meeting ended with a slightly different version of the 'simple or complex' query. We were asked to record our opinion, on a scale of one to ten, of the mechanisms of aging in wild-type mammals in age-structured populations. 'One' on this scale reflects the view that: 'There is only a single major mechanism of aging. This is modulated by only a few major gene effects.' A vote of 'ten' means that: 'There are numerous mechanisms of aging. These are modulated by hundreds or thousands of gene effects, both major and minor, depending upon the allele, gene-gene and gene-environmental interactions.' At the beginning of the meeting, the mean score was 7.3 with a standard error of 0.67; at the end of the meeting, the mean score was 7.5 with a standard error of 0.67. The genetics of aging is complicated, but not intractable.

Box 2

Linkage disequilibrium (LD) is technically a measure of the correlation in gene frequency between two polymorphic loci. Imagine that a new mutation affecting longevity occurs at some time in the dim and distant past. Initially, this mutant allele is associated (in LD with) all alleles present in the genome of the individual in which the mutation occurred. Over time, however, recombination between the original mutant and other polymorphic alleles restores 'linkage equilibrium': the observed two-locus genotype frequency is equal to the product of the genotype frequencies expected from each of the two loci separately, assuming random mating and no selection at either locus. Whether the gene frequencies at two loci are in LD thus depends on how recently the mutation occurred, the recombination distance between the loci, and the population size. In large randomly mating populations, only very closely linked loci will be in LD. Therefore, if one observes a difference in mean phenotype (length of life) between marker locus genotypes, one could infer that (in an ideal population satisfying all of the population genetic assumptions) the marker locus is closely linked to the polymorphism causing the life span effect. In other words, observing LD between a marker locus and a trait phenotype is exactly equivalent to QTL mapping, but on a much finer scale. LD mapping provides a means for fine-mapping QTLs, as well as evaluating the effects of candidate genes on quantitative traits, including longevity.